

# STOCHASTIC CHAPLYGIN SYSTEMS

SIMON HOCHGERNER

ABSTRACT. We mimic the stochastic Hamiltonian reduction of Lazaro-Cami and Ortega [17, 18] for the case of certain non-holonomic systems with symmetries.

Using the non-holonomic connection it is shown that the drift of the stochastically perturbed  $n$ -dimensional Chaplygin ball is a certain gradient of the density of the preserved measure of the deterministic system.

## CONTENTS

1. Introduction	1
2. Stochastic Chaplygin systems and reduction of symmetries	2
2.A. Stochastic geometry	2
2.B. Stochastic Hamiltonian systems	3
2.C. Stochastic $G$ -Chaplygin systems	4
3. The stochastic Chaplygin ball	5
3.A. The deterministic system	5
3.B. Brownian motion on the configuration space	6
3.C. Constrained Brownian motion and compression	7
4. Appendix: $G$ -Chaplygin systems and symmetry reduction	11
References	12

## 1. INTRODUCTION

Imagine a ball sitting on a rough horizontal table. Because of the roughness of the table this ball -the Chaplygin ball- cannot slip, but it can turn about the vertical axis without violating the constraints. Its geometric and gravitational center coincide but there may be an inhomogeneous mass distribution. Suppose now that the ball is subjected to a Brownian noise such that there is a random jiggling in all of its angular and translational degrees of freedom. This problem is similar to the stochastic rigid body considered in [18] but upon imposing the no-slip constraints some differences can be expected. One may wonder if the stochastic Chaplygin ball will acquire a drift that makes it roll on the table or spin about its vertical axis or both? The answer to this question is given by Theorem 3.3: The drift follows a Fick's law in the following sense. The configuration space of the  $n$ -dimensional ball is  $Q = \text{SO}(n) \times \mathbb{R}^{n-1}$  and there are  $n-1$  constraints corresponding to the directions in the table  $\mathbb{R}^{n-1}$ . By a symmetry reduction argument (compression) one can eliminate the  $\mathbb{R}^{n-1}$ -factor and the *deterministic* motion of the ball can be described by the geodesic equations of the so-called non-holonomic connection  $\nabla^{\text{nh}}$  on  $\text{SO}(n)$ . With respect to this connection one can now show that the process on  $\text{SO}(n)$  describing the balls *stochastic* motion is a non-holonomic diffusion separating into a drift- and a martingale-term. See Definition 3.2. Theorem 3.3 says that this  $\nabla^{\text{nh}}$ -drift equals

$$-\frac{1}{2}\text{grad}^{\mu_0} \log \mathcal{N}$$

where  $\mathcal{N}$  is the preserved density (3.5) of the deterministic ball and the gradient is computed with respect to the kinetic energy metric (3.4) of the ball.

In particular, when the ball is homogeneous there is no drift. Moreover, it is shown in Corollary 3.4 that the homogeneous ball's stochastic process factorizes to a Brownian motion on 'the ultimate reduced configuration space'  $S^{n-1} = (Q/\mathbb{R}^{n-1})/\mathrm{SO}(n-1)$ . This is in analogy to the corresponding deterministic case where the motion is Hamiltonian at the ultimate reduced level. (See [12].)

In 3 dimensions the drift  $-\frac{1}{2}\mathrm{grad}^{\mu_0} \log \mathcal{N}$  does not have an angular velocity component about the vertical axis of the ball in the space frame. For dimensions  $n > 3$  it turns out that this property is related to the Hamiltonization of the deterministic system: If the inertia matrix describing the balls mass distribution satisfies the Hamiltonization condition 3.7 then the drift does not have an angular velocity component about the vertical axis in the space frame. On the other hand, the drifts angular momentum about the vertical axis is always 0, regardless of the dimension or the mass distribution.

Section 2 starts by collecting some definitions and facts from stochastic differential geometry as presented in [14, 10]. Then we rehearse the basics of the stochastic Hamiltonian mechanics and their symmetries as introduced in [17, 18]. In Section 2.C these ideas are transferred to describe stochastic  $G$ -Chaplygin systems. It is noticed that the reduction of symmetries, termed *compression* in this context, works naturally.

In Section 3 this construction is applied to the  $n$ -dimensional Chaplygin ball. First some facts about the deterministic system, such as the preserved measure and Hamiltonization, are recalled. Then the Hamiltonian construction of [17] of Brownian motion on the configuration space  $Q = \mathrm{SO}(n) \times \mathbb{R}^{n-1}$  is reviewed. Section 3.C makes the constraint forces act on the Brownian motion according to the recipe of Section 2.C. Thus a constrained stochastic motion is obtained and the above mentioned Theorem 3.3 is found. Finally, we note that we can also treat the cases of angular jiggling only or horizontal jiggling only. The latter corresponds to the Chaplygin ball sitting on a table which undergoes a translational Brownian motion. The 3-dimensional version of this case has been considered in [19].

## 2. STOCHASTIC CHAPLYGIN SYSTEMS AND REDUCTION OF SYMMETRIES

**2.A. Stochastic geometry.** We state some notions from stochastic differential geometry. The references we used here are [10, 14]. See also the appendix of [17].

Let  $M$  be a manifold, let  $(\Omega, \mathcal{F}, \{\mathcal{F}_t : t \geq 0\}, P)$  be a filtered probability space, and let  $\Gamma : \mathbb{R}_+ \times \Omega \rightarrow M$  be an adapted stochastic process. (We consider only continuous processes.)

The process  $\Gamma$  is called a SEMI-MARTINGALE if  $f \circ \Gamma$  is a semi-martingale in  $\mathbb{R}$  for all  $f \in C^\infty(M)$ . See [10, Chapter III].

The definition of a martingale in  $M$  depends upon a choice of a connection. Let  $\nabla$  be a connection in  $TM \rightarrow M$ . Then the Hessian of  $\nabla$  is defined by

$$\mathrm{Hess}^\nabla(f)(X, Y) = XY(f) - \nabla_X Y(f)$$

for  $X, Y \in \mathfrak{X}(M)$ . This is bilinear in  $X$  and  $Y$  but not symmetric, in general, since  $\mathrm{Hess}^\nabla(f)(X, Y) - \mathrm{Hess}^\nabla(f)(Y, X) = -\mathrm{Tor}^\nabla(X, Y)(f)$ .

**Definition 2.1.** A semi-martingale  $\Gamma : \mathbb{R}_+ \times \Omega \rightarrow M$  is said to be a  $\nabla$ -MARTINGALE if, for any  $f \in C^\infty(M)$ ,

$$f \circ \Gamma_t - f \circ \Gamma_0 - \int_0^t \mathrm{Hess}^\nabla(f)(d\Gamma_s, d\Gamma_s) ds$$

is a local martingale in  $\mathbb{R}$ .

In [10, Chapter IV] this is stated in terms of a torsionless connection but it is noted that one can also allow for connections with torsion, since it is proved ([10, (3.14)]) that  $\int_0^t \mathrm{Hess}^\nabla(f)(d\Gamma_s, d\Gamma_s) ds$  depends only on the symmetric part of  $\mathrm{Hess}^\nabla$ . The situation is similar to the notion of a  $\nabla$ -geodesic. This also depends only on the torsionless part  $\nabla - \frac{1}{2}\mathrm{Tor}^\nabla$  of  $\nabla$ .

Let  $M$  and  $N$  be manifolds, let  $(\Omega, \mathcal{F}, \{\mathcal{F}_t : t \geq 0\}, P)$  be a filtered probability space, and let  $X : \mathbb{R}_+ \times \Omega \rightarrow N$  be a semi-martingale. A STRATONOVICH OPERATOR  $\mathcal{S}$  from  $TN$  to  $TM$  is a family of linear maps

$$\mathcal{S}_{(x,y)} : T_x N \longrightarrow T_y M$$

which depends smoothly on  $x \in N$  and  $y \in M$ . In other words,  $\mathcal{S}$  is a section of  $T^*N \otimes TM \rightarrow N \times M$ . A Stratonovich differential equation for a semi-martingale  $\Gamma : \mathbb{R}_+ \times \Omega \rightarrow M$  is written as

$$\delta\Gamma = \mathcal{S}(\Gamma, X)\delta X.$$

See [10, Chapter VII] for the precise meaning of this equation as well as existence and uniqueness (up to explosion time) of solutions.

Assume now that  $N = \mathbb{R} \times \mathbb{R}^n$  and that  $X : \mathbb{R}_+ \times \Omega \rightarrow \mathbb{R} \times \mathbb{R}^n$ ,  $(t, \omega) \mapsto (t, W_t(\omega))$  where  $W$  denotes  $n$ -dimensional Brownian motion. Let  $X_0, X_1, \dots, X_n$  be vectorfields on  $M$  and define the Stratonovich operator

$$\mathcal{S}(x, y) : \mathbb{R}^{n+1} \longrightarrow T_y M, (t, w^1, \dots, w^n) \longmapsto tX_0(y) + \sum w^i X_i(y).$$

Let  $f \in C^\infty(M)$ . Then  $f \circ \Gamma : \mathbb{R}_+ \times \Omega \rightarrow \mathbb{R}$  satisfies

$$\delta(f \circ \Gamma) = df(\Gamma) \cdot \mathcal{S}(X, \Gamma)\delta X = df(\Gamma) \cdot X_0(\Gamma)\delta t + \sum df(\Gamma) \cdot X_i(\Gamma)\delta W^i.$$

Hence  $\Gamma$  defines a diffusion in  $M$  by [14, Chapter V, Thm. 1.2]. The generator of this diffusion is the second order differential operator  $A$  given by

$$Af = X_0 f + \frac{1}{2} \sum X_i X_i f$$

where  $f \in C^\infty(M)$ .

**Definition 2.2.** If  $M$  is equipped with a Riemannian metric  $\mu$  and  $\Delta$  is the associated Laplacian then the diffusion is called BROWNIAN MOTION in  $(M, \mu)$  if

$$A = \frac{1}{2} \Delta.$$

This definition agrees with the one given in [10, (5.16)] or [20, (8.5.18)] but differs slightly from the one in [14, Chapter V, Def. 4.2] where it is required that  $n = \dim M$ .

**2.B. Stochastic Hamiltonian systems.** This section presents some of the concepts elaborated in [17, 18].

Let again  $N = \mathbb{R}^n$  and consider a Poisson manifold  $(M, \{.,.\})$  together with an  $\mathbb{R}^n$ -valued Hamiltonian function  $h = (h_1, \dots, h_n) : M \rightarrow \mathbb{R}^n$ . Let  $X : \mathbb{R}_+ \times \Omega \rightarrow \mathbb{R}^n$  be a semi-martingale. The associated stochastic Hamiltonian system is given by the Stratonovich equation  $\delta\Gamma = \mathcal{S}(X, \Gamma)\delta X$  where  $\mathcal{S}$  is defined in terms of the Hamiltonian structure, that is,

$$\mathcal{S}(x, y) : \mathbb{R}^n \longrightarrow T_y M, (x^1, \dots, x^n) \longmapsto \sum X_{h_i}(y) x^i$$

where the Hamiltonian vectorfield  $X_{h_i}$  is the vectorfield corresponding to the derivation  $\{h_i, .\}$ . When  $(M, \omega)$  is a symplectic manifold then one uses the Hamiltonian fields defined by  $i(X_{h_i})\omega = dh_i$ .

These systems allow for a symmetry reduction analogous to classical mechanics. We state the symplectic version of this theorem of [18, Section 6]:

**Theorem 2.3** (Stochastic Hamiltonian reduction). *Let  $(M, \omega)$  be a symplectic manifold with Hamiltonian  $h : M \rightarrow \mathbb{R}^n$  and stochastic component  $X$  as above. Assume that  $(M, \omega, h)$  are invariant under the free and proper action of a Lie group  $G$  such that a coadjoint equivariant momentum map  $J : M \rightarrow \mathfrak{g}^*$  exists. Fix a level  $\lambda \in \mathfrak{g}^*$ .*

*Then  $J^{-1}(\lambda)$  is invariant under the flow of  $\mathcal{S}$ . Moreover,  $\mathcal{S}$  induces a Stratonovich operator  $\mathcal{S}_\lambda$  from  $T\mathbb{R}^n$  to  $T(J^{-1}(\lambda)/G_\lambda)$  and solutions of  $\mathcal{S}$  with initial condition in  $J^{-1}(\lambda)$  project to solutions of  $\mathcal{S}_\lambda$ . The induced operator is given by*

$$\mathcal{S}_\lambda(x, y_0) : (x^1, \dots, x^n) \longmapsto \sum X_{h_i^\lambda}(y_0) x^i$$

*where  $x \in \mathbb{R}^n$ ,  $y_0 \in J^{-1}(\lambda)/G_\lambda$  and  $X_{h_i^\lambda}$  is the Hamiltonian vectorfield with respect to the reduced symplectic form on  $J^{-1}(\lambda)/G_\lambda$  of the induced function  $h_i^\lambda$ .*

**2.C. Stochastic  $G$ -Chaplygin systems.** A (deterministic)  $G$ -Chaplygin system consists of a Riemannian configuration manifold  $(Q, \mu)$ , a Lie group  $G$  acting freely and properly by isometries on  $(Q, \mu)$ , and a horizontal space  $\mathcal{D}$  of the principal bundle  $\pi : Q \rightarrow Q/G$ . Hence  $\mathcal{D}$  is the kernel of a connection form  $\mathcal{A} : TQ \rightarrow \mathfrak{g}$ . The Lagrangian of the system is the kinetic energy, i.e.,  $L(q, v) = \frac{1}{2}\mu_q(v, v)$ . In general,  $\mathcal{D}$  is *not*  $\mu$ -orthogonal to the vertical space  $\ker T\pi$ . We will henceforth identify  $TQ = T^*Q$  via  $\mu$ . See also Section 4.

Let  $J_G : TQ \rightarrow \mathfrak{g}^*$  denote the standard equivariant momentum map associated to the lifted  $G$ -action on  $TQ$ . Given a  $G$ -invariant function  $h : TQ \rightarrow \mathbb{R}$  one may use Noether's theorem to conclude that the Hamiltonian vectorfield  $X_h$  is tangent to  $J_G^{-1}(0)$  and, moreover, is projectable for  $J_G^{-1}(0) \rightarrow J_G^{-1}(0)/G = T(Q/G)$ . This is why results such as Theorem 2.3 work. Since  $\mathcal{D}$  is a perturbed version of  $J_G^{-1}(0)$  we look for a substitute construction.

Consider the horizontal space associated to the pulled back connection  $\tau^*\mathcal{A} = \mathcal{A} \circ T\tau : T(TQ) \rightarrow TQ \rightarrow \mathfrak{g}$  of the tangent lifted  $G$ -action on  $TQ$ ,

$$\mathcal{F} := \ker \tau^*\mathcal{A} \subset TTQ.$$

By assumption  $\mathcal{D}$  is also  $G$ -invariant. Thus we can consider the restricted  $G$ -action on  $\mathcal{D}$  and the associated connection  $\iota^*\tau^*\mathcal{A} : T\mathcal{D} \rightarrow \mathfrak{g}$  where  $\iota : \mathcal{D} \hookrightarrow TQ$  is the inclusion. Define

$$\mathcal{C} := \ker \iota^*\tau^*\mathcal{A} = \mathcal{F} \cap T\mathcal{D} \subset T\mathcal{D}$$

to be the horizontal space of the principal bundle  $\mathcal{D} \rightarrow \mathcal{D}/G = T(Q/G)$

According to [1] we can decompose  $TTQ$  along  $\mathcal{D}$  as

$$TTQ|_{\mathcal{D}} = \mathcal{C} \oplus \mathcal{C}^\Omega$$

where  $\mathcal{C}^\Omega$  is the  $\Omega$ -orthogonal of  $\mathcal{C}$  in  $TTQ|_{\mathcal{D}}$ . This implies that the fiber-wise restriction of  $\iota^*\Omega$  to  $\mathcal{C} \times \mathcal{C}$  is non-degenerate. For  $z \in \mathcal{D}$  define the projection

$$(2.1) \quad P_z : T_z(TQ) \longrightarrow \mathcal{C}_z$$

where we project along  $\mathcal{C}_z^\Omega$ . Moreover, for a  $k$ -form  $\phi$  on  $TQ$  we denote the fiber-wise restriction of  $\iota^*\phi$  to  $\Pi^k\mathcal{C}$  by  $\phi^{\mathcal{C}}$ . For a function  $h : TQ \rightarrow \mathbb{R}$  we may thus define the vectorfield  $X_h^{\mathcal{C}}$  on  $\mathcal{D}$  with values in  $\mathcal{C}$  by the formula

$$(2.2) \quad P_z X_h(z) = (\Omega^{\mathcal{C}})_z^{-1}(dh)_z^{\mathcal{C}} =: X_h^{\mathcal{C}}(z)$$

where  $z \in \mathcal{D}$ . The link to the non-holonomic system introduced above is the following: Via the Legendre transform the dynamics of the non-holonomic system  $(Q, \mathcal{D}, L)$  can be equivalently described by the triple  $(TQ, \Omega^{\mathcal{C}}, \mathcal{H} = \frac{1}{2}\mu(p, p))$  together with equation (2.2). Thus the Lagrange multipliers have been encoded in the two-form  $\Omega^{\mathcal{C}}$  or, equivalently, in the projection  $P : T(TQ)|_{\mathcal{D}} \rightarrow \mathcal{C}$ . The idea is that a non-holonomic system is a Hamiltonian system acted upon by constraint forces. The effect of the forces is described by the projector  $P : T(TQ)|_{\mathcal{D}} \rightarrow \mathcal{C}$ .

In addition to this structure consider now a semi-martingale  $X : \mathbb{R}_+ \times \Omega \rightarrow \mathbb{R}^n$  and a Hamiltonian function  $h = (h_i)_i : TQ \rightarrow \mathbb{R}^n$  as above. The associated stochastic non-holonomic system is given by the Stratonovich equation  $\delta\Gamma = \mathcal{S}^{\mathcal{C}}(X, \Gamma)\delta X$  where the Stratonovich operator  $\mathcal{S}^{\mathcal{C}} : \mathbb{R}^n \times \mathcal{D} \rightarrow \text{Hom}(T\mathbb{R}^n, \mathcal{C})$  is given by

$$(2.3) \quad \mathcal{S}^{\mathcal{C}}(n, z) : \mathbb{R}^n \longrightarrow \mathcal{C}_z, (x^1, \dots, x^n) \longmapsto \sum P_z X_{h_i}(z) x^i = \sum X_{h_i}^{\mathcal{C}}(z) x^i.$$

Thus the non-holonomic Stratonovich operator arises, by applying the constraint forces, as a projection of the Hamiltonian Stratonovich operator into  $\mathcal{C}$ . When the  $h^i$  are  $G$ -invariant we refer to the collection  $(Q, \mathcal{D}, h = (h_i)_i, X)$  as a stochastic  $G$ -Chaplygin system.

**Proposition 2.4** (Compression of stochastic  $G$ -Chaplygin systems). *Let  $(Q, \mathcal{D}, h = (h_i)_i, X)$  be a stochastic  $G$ -Chaplygin system with  $\Omega^{\mathcal{C}}$  as above. Then the Stratonovich operator (2.3) compresses to a Stratonovich operator  $\mathcal{S}^{\text{nh}}$  from  $T\mathbb{R}^n$  to  $\mathcal{D}/G = T(Q/G)$  which is given by*

$$\mathcal{S}^{\text{nh}}(n, z_0) : \mathbb{R}^n \longrightarrow T_{z_0}(T(Q/G)), (x^1, \dots, x^n) \longmapsto \sum X_{h_i^0}^{\text{nh}}(z) x^i$$

where  $z_0 \in T(Q/G)$  and  $h_i^0 : T(Q/G) \rightarrow \mathbb{R}$  is the function induced on the quotient from the invariant function  $\iota^*h_i$ . Moreover, solutions of (2.3) project to solutions of  $\mathcal{S}^{\text{nh}}$ .

*Proof.* Everything is entirely analogous to the proof of [18, Theorem 6.7] with the only difference that now one uses Proposition 4.1 instead of the usual symplectic reduction theorem.  $\square$

We think of  $\delta\Gamma = \mathcal{S}^{\text{nh}}(X, \Gamma)\delta X$  as the equations of motion of the system  $(Q, \mathcal{D}, h, X)$ .

**Proposition 2.5** (Ito representation). *Let  $f \in C^\infty(T(Q/G))$ . Then the Ito representation of the equation  $\delta(f \circ \Gamma) = \mathcal{S}^{\text{nh}}(X, f \circ \Gamma)\delta X$  is*

$$d(f \circ \Gamma) = \sum_i X_{h_i^0}^{\text{nh}} f(\Gamma) dX^i + \frac{1}{2} \sum_{i,j} X_{h_i^0}^{\text{nh}} X_{h_j^0}^{\text{nh}} f(\Gamma) [dX^j, dX^i]$$

where  $X^i = \text{pr}^i \circ X$ .

*Proof.* This follows exactly as in the proof [17, Proposition 2.3]. It is only necessary to notice that this proof does not depend on whether or not the non-holonomic bracket  $X_g^{\text{nh}} f = \{g, f\}^{\text{nh}} = -\{f, g\}^{\text{nh}}$  satisfies the Jacobi identity. The only property of the Poisson bracket which is used in [17, Proposition 2.3] is the Leibniz rule and this feature is evidently shared by the non-holonomic bracket.  $\square$

### 3. THE STOCHASTIC CHAPLYGIN BALL

**3.A. The deterministic system.** For background on the Chaplygin ball we refer to [6, 8, 9, 11, 12, 16]. The configuration space of Chaplygin's  $n$ -dimensional rolling ball is  $Q = K \times V$  where  $K = \text{SO}(n)$  and  $V = \mathbb{R}^{n-1}$ . The no-slip constraints are given by the distribution  $\mathcal{D} = (\mathcal{A} + \text{pr}_2)^{-1}(0) \subset TQ$  where

$$\mathcal{A} = \sum \eta^a \otimes e_a$$

where  $e_1, \dots, e_{n-1}$  is the standard basis on  $V$  and we stick to the following conventions:<sup>1</sup>  $TK = K \times \mathfrak{k}$  is trivialized via left-multiplication and  $\mathfrak{k}$  is equipped with the Ad-invariant inner product  $\langle \cdot, \cdot \rangle$ .<sup>2</sup> Let  $H = \text{SO}(n-1) \subset K$  be the stabilizer in  $K$  of the  $n$ -th standard vector  $e_n$  such that  $H$  acts on  $V$  in the natural way. We decompose  $\mathfrak{k} = \mathfrak{h} \oplus \mathfrak{h}^\perp$  with respect to the Ad-invariant inner product  $\langle \cdot, \cdot \rangle$  on  $\mathfrak{k}$ . With respect to this inner product we introduce an orthonormal system

$$Y_\alpha, \alpha = 1, \dots, k = \dim \mathfrak{h} \text{ and } Z_a, a = 1, \dots, n-1$$

on  $\mathfrak{k}$  such that  $Y_\alpha \in \mathfrak{h}$  and  $Z_a \in \mathfrak{h}^\perp$ . Associated to this basis we define the right invariant vector fields  $\xi_\alpha$  and  $\zeta_a$ . In the left trivialization these read

$$\xi_\alpha(s) = \text{Ad}(s^{-1})Y_\alpha \text{ and } \zeta_a(s) = \text{Ad}(s^{-1})Z_a.$$

Dually we introduce the corresponding right invariant coframe

$$\rho^\alpha = \langle \xi_\alpha, \cdot \rangle \text{ and } \eta^a = \langle \zeta_a, \cdot \rangle.$$

The Lagrangian is the function

$$L : TQ = K \times \mathfrak{k} \times TV \longrightarrow \mathbb{R}, (s, u, x, x') \longmapsto \frac{1}{2} \langle \mathbb{I}u, u \rangle + \frac{1}{2} \langle x', x' \rangle.$$

where  $\mathbb{I}$  is the inertia matrix in body coordinates. The rolling ball with the no-slip constraint is the non-holonomic system described by the data  $(Q, \mathcal{D}, L)$  where the equations of motion follow from the Lagrange-d'Alembert principle. However, we will not have much use for the Lagrange function below since we will only perturb the resting ball. Note also that we overload the symbol  $\langle \cdot, \cdot \rangle$  by using it for the Euclidean inner product on  $V$  as well as for the Ad-invariant structure on  $\mathfrak{k}$ .

From a structural point of view the decisive feature of the Chaplygin ball is that its constraints are given by a connection  $\text{pr}_2 + \mathcal{A} : TQ \rightarrow V$  on the (trivial) principal bundle  $V \hookrightarrow K \times V \rightarrow K$  where  $V$  acts on itself by addition. Thus  $\mathcal{D}$  is the horizontal space of this connection. However,  $\mathcal{D}$  is not  $\mu$ -orthogonal to the vertical space of the bundle. The fact that the system is non-holonomic is reflected in the non-flatness,  $\text{Curv}^{\mathcal{A}} = d\mathcal{A} \neq 0$ .

<sup>1</sup>In Section 2.C it was the connection form  $\mathcal{A} + \text{pr}_2$  which was called  $\mathcal{A}$ .

<sup>2</sup>The space  $\mathfrak{X}_L(K)$  of left invariant vectorfields and the Lie algebra  $\mathfrak{k}$  will also be identified without further notice.

Compression of the Chaplygin ball system yields the almost Hamiltonian system  $(TK, \Omega_{\text{nh}}, \mathcal{H}_c)$  where  $\Omega_{\text{nh}}$  is described in Proposition 4.1,  $TK$  and  $T^*K$  are identified via the induced metric

$$(3.4) \quad \mu_0(u_1, u_2)_s = \langle \mathbb{I}u_1, u_2 \rangle + \langle \mathcal{A}_s(u_1), \mathcal{A}_s(u_2) \rangle = \langle (\mathbb{I} + \mathcal{A}_s^* \mathcal{A}_s)u_1, u_2 \rangle,$$

and  $\mathcal{H}_c = \frac{1}{2}\mu_0(u, u)_s$ . The metric  $\mu_0$  is the sum of a left invariant and a right invariant term. Thus it constitutes an  $L + R$ -system, see [15]. Note the useful formula  $\mathcal{A}^* \mathcal{A}(u) = \sum \langle \zeta_a, u \rangle \zeta_a$ .

The compressed system  $(TK, \Omega_{\text{nh}}, \mathcal{H}_c)$  is further invariant under the lift of the left multiplication action of  $H$  on  $K$ . Physically this corresponds to rotation of the ball about the  $e_n$ -axis in the space frame. This is an *inner symmetry* and gives, by the non-holonomic Noether theorem, rise to a conserved quantity. This quantity is just the standard momentum map  $J_H : TK \rightarrow \mathfrak{h}^* = \mathfrak{h}$ ,  $(s, u) \mapsto \sum \rho_s^\alpha (\mathbb{I}u) Y_\alpha$  of the (co-)tangent lifted  $H$ -action on  $TK$ .

The Chaplygin ball shares an important feature with Hamiltonian systems. Namely, it possesses a preserved measure ([6, 11]). At the compressed level -the  $TK$ -level- the density  $\mathcal{N} : K \rightarrow \mathbb{R}$  of this measure with respect to the Liouville volume on  $TK$  is

$$(3.5) \quad \mathcal{N}(s) = (\det \mu_0(s))^{-\frac{1}{2}}.$$

This function plays the central role in all questions of Hamiltonization of the system. Note that  $\mathcal{N}$  is  $H$ -invariant and thus descends to a function  $\mathcal{N} : K/H = S^{n-1} \rightarrow \mathbb{R}$ . For further reference we also record that

$$(3.6) \quad d(\log \mathcal{N}) = \sum \langle [\mu_0^{-1} \zeta_a, \zeta_a], \zeta_b \rangle \eta^b.$$

In [12] it is proved that  $\Omega_{\text{nh}}$  can be replaced by  $\tilde{\Omega} = \Omega^K - \frac{1}{2} \sum \langle [\zeta_a, \zeta_b], \cdot \rangle \eta^a \wedge \eta^b$  without altering the equations of motion, that is,  $i(X_{\mathcal{H}_c}^{\text{nh}}) \Omega_{\text{nh}} = i(X_{\mathcal{H}_c}^{\text{nh}}) \tilde{\Omega} = d\mathcal{H}_c$ . Now the new system  $(TK, \tilde{\Omega}, \mathcal{H}_c)$  has the same dynamics but has the advantage of being liable to reduction with respect to the internal symmetry group  $H$ : Let  $J_H : TK \rightarrow \mathfrak{h}^*$  be the momentum map introduced above,  $\lambda \in \mathfrak{h}^*$  and  $\iota : J_H^{-1}(\lambda) \hookrightarrow TK$  the inclusion. Then  $\iota^* \tilde{\Omega}$  descends to an almost symplectic two form  $\tilde{\Omega}_\lambda$  on  $J_H^{-1}(\lambda)/H_\lambda$ . ( $\iota^* \Omega_{\text{nh}}$  is not horizontal for the projection onto  $J_H^{-1}(\lambda)/H_\lambda$ .) In this way one can recover the Hamiltonization of the 3-dimensional ball of [4, 5]: It is shown in [12, Proposition 4.4] that  $d(\mathcal{N} \tilde{\Omega}_\lambda) = 0$  if  $n = 3$ . Thus, for  $n = 3$ , the rescaled vectorfield  $\mathcal{N}^{-1} X_{\mathcal{H}_c}^{\text{nh}}$  is Hamiltonian with respect to  $(J_H^{-1}(\lambda)/H_\lambda = TS^2, \mathcal{N} \tilde{\Omega}_\lambda, \mathcal{H}_c)$ . Moreover, the homogeneous ball,  $\mathbb{I} = 1$ , is Hamiltonian at the  $J_H^{-1}(\lambda)/H_\lambda$ -level for any dimension  $n$ . It is interesting to notice that none of these statements hold at the  $TK$ -level.

In [13, Section 4.A] the following Hamiltonization condition is proved for arbitrary dimension  $n$ : Let  $\lambda = 0 \in \mathfrak{h}^*$ . Then  $d(\mathcal{N} \tilde{\Omega}_0) = 0$  if and only if

$$(3.7) \quad (n-2) \langle \mu_0^{-1} \zeta_d, [\zeta_b, \zeta_c] \rangle = \sum_a \langle \mu_0^{-1} \zeta_a, [\zeta_b, \zeta_a] \delta_{c,d} - [\zeta_c, \zeta_a] \delta_{b,d} \rangle$$

which is an algebraic condition on  $\mathbb{I}$ . For  $n = 3$  this condition is trivially satisfied. In the stochastic context this condition appears in Theorem 3.3 below.

**3.B. Brownian motion on the configuration space.** We follow [17] to construct Brownian motion on  $Q = K \times V$ . Let  $\nabla^\mu$  be the Levi-Civita connection of

$$\mu((u, x'), (v, y')) = \langle \mathbb{I}u, v \rangle + \langle x', y' \rangle.$$

Thus for  $u, v \in \mathfrak{k}$  we have

$$\nabla_u^\mu v = (\nabla_u^\mathbb{I} v, 0) = (\frac{1}{2}[u, v] + \frac{1}{2}\mathbb{I}^{-1}([u, \mathbb{I}v] + [v, \mathbb{I}u]), 0)$$

where  $\nabla^\mathbb{I}$  is the Levi-Civita connection of the left-invariant metric defined by  $\langle \mathbb{I}, \cdot \rangle$  on  $K$ . Note that we identify  $\mathfrak{k} \cong \mathfrak{k}^*$  via  $\langle \cdot, \cdot \rangle$  whence  $\mathbb{I} : \mathfrak{k} \rightarrow \mathfrak{k}^* \cong \mathfrak{k}$ . Let

$$v_1, \dots, v_m, m = \dim \mathfrak{k} = \frac{1}{2}n(n-1)$$

denote a basis which is orthonormal for  $\langle \mathbb{I}, \cdot \rangle$ . Thus  $v_1, \dots, v_m, e_1, \dots, e_{n-1}$  is a left invariant frame on  $Q$  which is orthonormal with respect to the left invariant metric  $\mu$ . (Remember that we identify  $TQ$  and  $T^*Q$

via  $\mu$ .) Let the functions  $H_0, H_i, F_a : TQ \rightarrow \mathbb{R}$  be given by

$$(3.8) \quad H_0(s, u, x, x') = -\frac{1}{2}\langle \mathbb{I}u, \sum \nabla_{v_i}^{\mathbb{I}} v_i \rangle, \quad H_i(s, u, x, x') = \langle \mathbb{I}u, v_i \rangle, \quad \text{and} \quad F_a(s, u, x, x') = \langle x', e_a \rangle.$$

Consider the semi-martingale

$$X : \mathbb{R}_+ \times \Omega \longrightarrow \mathbb{R} \times \mathbb{R}^m \times \mathbb{R}^{n-1}, \quad (t, \omega) \longmapsto (t, B_t^1(\omega), \dots, B_t^m(\omega), W_t^1(\omega), \dots, W_t^{n-1}(\omega))$$

where  $B^i, W^a$  are  $m+n-1$  independent Brownian motions. The Stratonovich stochastic differential equation which is associated to these data is

$$(3.9) \quad \delta\Gamma = \mathcal{S}^{\text{Ham}}(X, \Gamma)\delta X$$

where the Stratonovich operator from  $T(\mathbb{R} \times \mathbb{R}^m \times \mathbb{R}^{n-1})$  to  $TK \times TV$  is defined by

$$\mathcal{S}^{\text{Ham}}(n; s, u, x, x')(t, b^1, \dots, b^m, w^1, \dots, w^{n-1}) = X_{H_0}(s, u, x, x')t + \sum X_{H_i}(s, u, x, x')b^i + \sum X_{F_a}(s, u, x, x')w^a$$

with  $X_H$  denoting the canonical Hamiltonian vectorfield of a function  $H : TQ \rightarrow \mathbb{R}$ . Using the Ito representation of this equation [17] show that the solutions  $\Gamma$  project via  $\tau : TQ \rightarrow Q$  onto Brownian motion on  $Q$ .

Using again the setting of [17] and Theorem 2.3 it is easy to see the following. Consider the  $V$ -action on  $Q$  as above. Let  $J_V^{-1}(0) = \{(s, u, x, 0)\}$  be the 0-level set of the standard momentum map  $J_V : TQ \rightarrow V^* = V$  of the lifted  $V$ -action on  $TQ$ . Then the Stratonovich equation (3.9) induces a Stratonovich equation  $\delta\Gamma_0 = \mathcal{S}_0(X, \Gamma_0)\delta X$  on  $J_V^{-1}(0)/V = TK$  and solutions  $\Gamma$  with initial condition in  $J_V^{-1}(0)$  project onto solutions  $\Gamma_0$ . Moreover,  $\tau_K \circ \Gamma_0$  is a Brownian motion on  $K$  where  $\tau_K : TK \rightarrow K$ .

When we regard  $\mathcal{D}$  as perturbed version of  $J_V^{-1}(0)$  we can ask how much of this observation remains true? This is the content of Section 3.C.

**3.C. Constrained Brownian motion and compression.** We now force the Brownian motion on  $Q$  to satisfy the constraints induced by  $\mathcal{D}$ . In accordance with Section 2.C we do so by applying the constraint forces to the Stratonovich operator from (3.9). Thus we are concerned with the equation  $\delta\Gamma = \mathcal{S}^{\text{nh}}(X, \Gamma)\delta X$  where  $X$  and  $H_0, H_i, F_a$  are as above and

$$\mathcal{S}^{\text{C}}(n, z) = P_z \mathcal{S}^{\text{Ham}}(n, z) : \mathbb{R} \times \mathbb{R}^m \times \mathbb{R}^{n-1} \longrightarrow \mathcal{C}_z \subset T_z \mathcal{D}$$

where  $P$  was defined in (2.1) and  $z = (s, u, x, -\mathcal{A}_s(u)) \in \mathcal{D}$ . The functions  $H_0, H_i, F_a$  are  $V$ -invariant and compress to functions  $h_0, h_i, f_a$  given by

$$\begin{aligned} h_0(s, u) &= H_0(s, u, x, -\mathcal{A}_s(u)) = -\frac{1}{2}\mu((u, -\mathcal{A}_s(u)), \sum \nabla_{v_i}^{\mu} v_i) \\ &= -\frac{1}{2}\langle \mathbb{I}u, \sum \nabla_{v_i}^{\mathbb{I}} v_i \rangle = -\frac{1}{2}\mu_0(u, \mu_0^{-1}\mathbb{I}\sum \nabla_{v_i}^{\mathbb{I}} v_i), \\ h_i(s, u) &= \langle \mathbb{I}u, v_i \rangle = \mu_0(u, \mu_0^{-1}\mathbb{I}v_i), \\ f_a(s, u) &= -\langle \mathcal{A}_s(u), e_a \rangle = -\mu_0(u, \mu_0^{-1}\zeta_a(s)). \end{aligned}$$

Notice that  $h_0$  and  $h_i$  are left invariant while the  $f_a$  are right invariant. The compressed non-holonomic Stratonovich operator is now of the form

$$\mathcal{S}^{\text{nh}}(n; s, u)(t, b^1, \dots, b^m, w^1, \dots, w^{n-1}) = X_{h_0}^{\text{nh}}(s, u)t + \sum X_{h_i}^{\text{nh}}(s, u)b^i + \sum X_{f_a}^{\text{nh}}(s, u)w^a.$$

We think of solutions of  $\delta\Gamma = \mathcal{S}^{\text{nh}}(X, \Gamma)\delta X$  as non-holonomic diffusions. This is in analogy to [2, Chapter V] where Hamiltonian diffusions are considered in a similar manner.

For a function  $f \in C^\infty(K)$ , viewed as a function on  $TK$  via pull-back, we have

$$(3.10) \quad X_{h_0}^{\text{nh}}(f) = -\frac{1}{2}df \cdot \mu_0^{-1}\mathbb{I}\sum \nabla_{v_i}^{\mathbb{I}} v_i, \quad X_{h_i}^{\text{nh}}(f) = df \cdot \mu_0^{-1}\mathbb{I}v_i, \quad \text{and} \quad X_{f_a}^{\text{nh}}(f) = -df \cdot \mu_0^{-1}\zeta_a.$$

Let  $\Gamma$  be the solution semi-martingale to  $\delta\Gamma = \mathcal{S}^{\text{nh}}(X, \Gamma)\delta X$  and let  $\tau : TK \rightarrow K$  be the projection. Then  $\tau \circ \Gamma$  solves

$$\delta(\tau \circ \Gamma) = T\tau \cdot \delta\Gamma = -\frac{1}{2}\mu_0^{-1}\mathbb{I}\sum \nabla_{v_i}^{\mathbb{I}} v_i(\tau \circ \Gamma)\delta t + \sum \mu_0^{-1}\mathbb{I}v_i(\tau \circ \Gamma)\delta B^i - \sum \mu_0^{-1}\zeta_a(\tau \circ \Gamma)\delta W^a.$$

According to [14, Chapter V, Theorem 1.2] this means that the semi-martingale  $\tau \circ \Gamma$  defines a diffusion in  $K$  whose generator is the second order differential operator

$$(3.11) \quad -\frac{1}{2}\mu_0^{-1}\mathbb{I}\sum\nabla_{v_i}^{\mathbb{I}}v_i + \frac{1}{2}\sum(\mu_0^{-1}\mathbb{I}v_i)(\mu_0^{-1}\mathbb{I}v_i) + \frac{1}{2}\sum(\mu_0^{-1}\zeta_a)(\mu_0^{-1}\zeta_a).$$

To identify the drift of the diffusion  $\tau \circ \Gamma$  a connection is needed. We introduce the non-holonomic connection which is explained in [7, Section 5.1.1].<sup>3</sup> Let  $\text{Pr}^\mu : TQ \rightarrow \mathcal{D}$  denote the projection onto  $\mathcal{D}$  along the  $\mu$ -orthogonal  $\mathcal{D}^\mu$  of  $\mathcal{D}$ . Note that  $\mathcal{D}^\mu \neq \ker T\pi$  where  $\pi : Q \rightarrow K$  is the projection. Let  $\text{hl}^A : \mathfrak{X}(K) \rightarrow \mathfrak{X}(Q; \mathcal{D})$ ,  $\text{hl}^A(s, x)(u) = (s, u, x, -\mathcal{A}_s(u))$  be the horizontal lift map associated to  $\mathcal{A}$ . Given  $X, Y \in \mathfrak{X}(K)$  the non-holonomic connection is prescribed by

$$(3.12) \quad \nabla_X^{\text{nh}}Y = T\pi\text{Pr}^\mu\nabla_{\text{hl}^AX}^\mu(\text{hl}^AY).$$

This connection is metric, i.e.,  $\nabla^{\text{nh}}\mu_0 = 0$ , and its geodesic equations are exactly the equations of motion of the non-holonomic system described by  $(TK, \Omega_{\text{nh}}, \mathcal{H}_c)$ . However, in general,  $\nabla^{\text{nh}}$  will have non-trivial torsion.

**Lemma 3.1.** *For  $u, v \in \mathfrak{k} = \mathfrak{X}_L(K)$  we have*

$$\nabla_u^{\text{nh}}v = \mu_0^{-1}(\mathbb{I}\nabla_u^{\mathbb{I}}v + \mathcal{A}^*\mathcal{A}[u, v])$$

where  $\nabla_u^{\mathbb{I}}v = \frac{1}{2}[u, v] + \frac{1}{2}\mathbb{I}^{-1}([u, \mathbb{I}v] + [v, \mathbb{I}u])$ . Its torsion is given by  $\text{Tor}^{\text{nh}}(u, v) = \mu_0^{-1}\mathcal{A}^*\mathcal{A}[u, v]$ .

Note that  $\nabla_u^{\text{nh}}v$  is not left invariant any more. At the compressed level the equations of motion of the Chaplygin ball write as  $u' + \mu_0^{-1}[u, \mathbb{I}u] = 0$ . In 3D this corresponds to [8, Equation (3.5)].

*Proof.* For  $u, v \in \mathfrak{k}$  we need to compute

$$\text{Pr}^\mu\nabla_{(u, -\mathcal{A}u)}^\mu(v, -\mathcal{A}v) = \text{Pr}^\mu(\nabla_u^{\mathbb{I}}v - u.\mathcal{A}v) = \text{Pr}^\mu(\nabla_u^{\mathbb{I}}v - \mathcal{A}[u, v]).$$

Now note that  $(w, X) \in \mathcal{D}^\mu$  if and only if  $w = \mathbb{I}^{-1}\mathcal{A}^*X$ . We have to solve

$$0 = \mathcal{A}(\nabla_u^{\mathbb{I}}v + \mathbb{I}^{-1}\mathcal{A}^*X) - \mathcal{A}[u, v] + X$$

for  $\mathbb{I}^{-1}\mathcal{A}^*X$ . The solution is found to be given by  $\mu_0\mathbb{I}^{-1}\mathcal{A}^*X = \mathcal{A}^*\mathcal{A}([u, v] - \nabla_u^{\mathbb{I}}v)$ . Therefore,

$$(3.13) \quad \nabla_u^{\text{nh}}v = \nabla_u^{\mathbb{I}}v + \mu_0^{-1}\mathcal{A}^*\mathcal{A}([u, v] - \nabla_u^{\mathbb{I}}v) = \mu_0^{-1}\mathcal{A}^*\mathcal{A}[u, v] + \mu_0^{-1}\mathbb{I}\nabla_u^{\mathbb{I}}v.$$

□

Recall that the Hessian of  $\nabla^{\text{nh}}$  is defined by

$$\text{Hess}^{\text{nh}}(f)(X, Y) = XY(f) - \nabla_X^{\text{nh}}Y(f)$$

for  $X, Y \in \mathfrak{X}(K)$  and  $f \in C^\infty(K)$ .

Let  $f \in C^\infty(K)$  and  $\delta\Gamma = \mathcal{S}^{\text{nh}}(X, \Gamma)\delta X$ . By Proposition 2.5, equations (3.10) and  $[dB^i, dB^j] = \delta^{i,j}dt$ ,  $[dW^a, dW^b] = \delta^{a,b}dt$ , as well as  $[dB^i, dW^a] = 0$  we have the Ito equation

$$\begin{aligned} d(f \circ \Gamma) &= \left( -\frac{1}{2}\mu_0^{-1}\mathbb{I}\sum\nabla_{v_i}^{\mathbb{I}}v_i + \frac{1}{2}\sum(\mu_0^{-1}\mathbb{I}v_i)(\mu_0^{-1}\mathbb{I}v_i) + \frac{1}{2}\sum(\mu_0^{-1}\zeta_a)(\mu_0^{-1}\zeta_a) \right) f(\Gamma)dt \\ &\quad + \sum(\mu_0^{-1}\mathbb{I}v_i)f(\Gamma)dB^i - \sum(\mu_0^{-1}\zeta_a)f(\Gamma)dW^a \\ &= \left( -\frac{1}{2}\mu_0^{-1}\mathbb{I}\sum\nabla_{v_i}^{\mathbb{I}}v_i + \frac{1}{2}\sum\nabla_{\mu_0^{-1}\mathbb{I}v_i}^{\text{nh}}(\mu_0^{-1}\mathbb{I}v_i) + \frac{1}{2}\sum\nabla_{\mu_0^{-1}\zeta_a}^{\text{nh}}(\mu_0^{-1}\zeta_a) \right) \\ &\quad + \frac{1}{2}\sum\text{Hess}^{\text{nh}}(\mu_0^{-1}\mathbb{I}v_i, \mu_0^{-1}\mathbb{I}v_i) + \frac{1}{2}\sum\text{Hess}^{\text{nh}}(\mu_0^{-1}\zeta_a, \mu_0^{-1}\zeta_a) \Big) f(\Gamma)dt \\ &\quad + \sum(\mu_0^{-1}\mathbb{I}v_i)f(\Gamma)dB^i - \sum(\mu_0^{-1}\zeta_a)f(\Gamma)dW^a, \end{aligned}$$

which also confirms (3.11). Having split the generator into first and purely second order part it makes sense to say what we mean by drift.

**Definition 3.2.** The vectorfield

$$-\frac{1}{2}\mu_0^{-1}\mathbb{I}\sum\nabla_{v_i}^{\mathbb{I}}v_i + \frac{1}{2}\sum\nabla_{\mu_0^{-1}\mathbb{I}v_i}^{\text{nh}}(\mu_0^{-1}\mathbb{I}v_i) + \frac{1}{2}\sum\nabla_{\mu_0^{-1}\zeta_a}^{\text{nh}}(\mu_0^{-1}\zeta_a)$$

is called the DRIFT of the diffusion  $\tau \circ \Gamma$  with respect to  $\nabla^{\text{nh}}$ .

<sup>3</sup>Contrary to [7] we only use the projected version of the non-holonomic connection.

Note that  $\tau \circ \Gamma$  is a  $\nabla^{\text{nh}}$ -martingale if and only if the  $\nabla^{\text{nh}}$ -drift vanishes. See also [10, Theorem (7.31)].

**Theorem 3.3.** *Let  $\Gamma$  be a solution of the Stratonovich equation  $\delta\Gamma = \mathcal{S}^{\text{nh}}(X, \Gamma)\delta X$  and let  $\tau : TK \rightarrow K$  be the projection.*

- (1) *Then, with respect to the non-holonomic connection  $\nabla^{\text{nh}}$  introduced in (3.12), the semi-martingale  $\tau \circ \Gamma$  defines a diffusion on  $K$  whose drift is the gradient  $-\frac{1}{2}\text{grad}^{\mu_0}(\log \mathcal{N})$  where  $\mathcal{N}$  is the density function defined in (3.5).*
- (2) *The drift  $-\frac{1}{2}\text{grad}^{\mu_0}(\log \mathcal{N})$  is horizontal with respect to the mechanical connection,  $\text{Hor}^{\text{mech}} = (\ker T\kappa)^{\mu_0^\perp}$ , on the principal bundle  $\kappa : K \rightarrow K/H = S^{n-1}$ . If  $\mathbb{I}$  satisfies the Hamiltonization condition (3.7) then the drift is also horizontal with respect to the principal bundle connection  $\sum \rho^\alpha \otimes Y_\alpha : TK \rightarrow \mathfrak{h}$  on  $\kappa : K \rightarrow K/H$ .*

Item (2) means that the drift's component of angular momentum about the vertical axis in the space frame vanishes, and when the Hamiltonization condition holds then the same is true for the component of angular velocity about the vertical axis. For  $n = 3$  this condition is always satisfied.

**Corollary 3.4.** *Let  $\Gamma$  be a solution of the Stratonovich equation  $\delta\Gamma = \mathcal{S}^{\text{nh}}(X, \Gamma)\delta X$  and let  $\tau : TK \rightarrow K$  and  $\kappa : K \rightarrow K/H = S^{n-1}$  be the obvious projections. Suppose that  $\mathbb{I} = 1$  (i.e., the ball is homogeneous).*

- (1) *Then  $\tau \circ \Gamma$  defines a martingale in  $K$  with respect to the non-holonomic connection.*
- (2) *The process  $\kappa \circ \tau \circ \Gamma$  is a Brownian motion on  $S^{n-1}$  whose generator is  $\frac{1}{2}$  times the Laplacian of  $\nu$ , where  $\nu$  is the metric on  $S^{n-1}$  induced from the left  $H$ -invariant metric  $\mu_0 = \langle (1 + \mathcal{A}^*\mathcal{A}), \cdot \rangle$  on  $K$ .*

Note that the restriction of  $\mu_0$  to  $\text{Hor} = \text{Hor}^{\text{mech}} = \text{span}\{\zeta_a\}$  equals twice the restriction of the biinvariant metric. This corollary is intuitive but nevertheless not obvious since the dynamics at the compressed level can never be described by a Hamiltonian reduction procedure. This is because  $\mathcal{A}$  is not the mechanical connection, even if the ball is homogeneous. (Compare with [12, Corollary 4.3].) Thus it does not fall in the category of [17, 18].

For reference we note the formula

$$(3.14) \quad v \cdot \mu_0(u) = v \cdot \sum \langle \zeta_a, u \rangle \zeta_a = \sum \langle \zeta_a, [v, u] \rangle \zeta_a - \sum \langle \zeta_a, u \rangle [v, \zeta_a] = \mathcal{A}^* \mathcal{A}[v, u] - [v, \mathcal{A}^* \mathcal{A}u].$$

*Proof of Theorem 3.3.* Let  $f \in C^\infty(K)$  which we regard via pull-back as a function on  $TK$ . According to Definition 3.2 we need to show that

$$(3.15) \quad -\mu_0^{-1} \mathbb{I} \sum \nabla_{v_i}^{\mathbb{I}} v_i + \sum \nabla_{\mu_0^{-1} \mathbb{I} v_i}^{\text{nh}} (\mu_0^{-1} \mathbb{I} v_i) + \sum \nabla_{\mu_0^{-1} \zeta_a}^{\text{nh}} (\mu_0^{-1} \zeta_a) = -\text{grad}^{\mu_0}(\log \mathcal{N}).$$

Claim:

$$(3.16) \quad \sum \nabla_{\mu_0^{-1} \mathbb{I} v_i}^{\text{nh}} (\mu_0^{-1} \mathbb{I} v_i) = \mu_0^{-1} \mathbb{I} \sum \nabla_{v_i}^{\mathbb{I}} v_i - \text{grad}^{\mu_0}(\log \mathcal{N}).$$

Indeed, we use (3.14) and the fact that  $u = \sum \langle \mathbb{I} v_i, u \rangle v_i$  to see that

$$\begin{aligned} \sum \nabla_{\mu_0^{-1} \mathbb{I} v_i}^{\text{nh}} (\mu_0^{-1} \mathbb{I} v_i) &= \sum \langle \mathbb{I} \mu_0^{-1} \mathbb{I} v_i, v_j \rangle \langle \mathbb{I} \mu_0^{-1} \mathbb{I} v_i, v_k \rangle \nabla_{v_j}^{\text{nh}} v_k + \sum \langle \mathbb{I} \mu_0^{-1} \mathbb{I} v_i, v_j \rangle (v_j \cdot \langle \mathbb{I} \mu_0^{-1} \mathbb{I} v_i, v_k \rangle) v_k \\ &= \sum \mu_0^{-1} [\mu_0^{-1} \mathbb{I} v_i, \mathbb{I} \mu_0^{-1} \mathbb{I} v_i] + \sum \langle \mathbb{I} \mu_0^{-1} \mathbb{I} v_i, v_j \rangle \langle \mathbb{I} \mu_0^{-1} [v_j, \mathcal{A}^* \mathcal{A} \mu_0^{-1} \mathbb{I} v_i], v_k \rangle v_k \\ &= \sum \mu_0^{-1} [\mu_0^{-1} \mathbb{I} v_i, \mathbb{I} \mu_0^{-1} \mathbb{I} v_i] + \sum \mu_0^{-1} [\mu_0^{-1} \mathbb{I} v_i, \mathcal{A}^* \mathcal{A} \mu_0^{-1} \mathbb{I} v_i] \\ &= \sum \mu_0^{-1} [\mu_0^{-1} \mathbb{I} v_i, \mathbb{I} v_i] \\ &= \sum \mu_0^{-1} [v_i, \mathbb{I} v_i] - \sum \mu_0^{-1} [\mu_0^{-1} \mathcal{A}^* \mathcal{A} v_i, \mathbb{I} v_i] \end{aligned}$$

where we use  $\mu_0^{-1} \mathbb{I} = 1 - \mu_0^{-1} \mathcal{A}^* \mathcal{A}$  in the last equation. Notice that  $[v_i, \mathbb{I} v_i] = \mathbb{I} \nabla_{v_i}^{\mathbb{I}} v_i$ . For the gradient part of claim (3.16) we consider

$$\begin{aligned} \langle \sum [\mu_0^{-1} \mathcal{A}^* \mathcal{A} v_i, \mathbb{I} v_i], \zeta_b \rangle &= -\sum \langle \mathbb{I} v_i, [\mu_0^{-1} \langle \zeta_a, v_i \rangle \zeta_a, \zeta_b] \rangle = -\sum \langle \zeta_a, \langle \mathbb{I} v_i, [\mu_0^{-1} \zeta_a, \zeta_b] \rangle v_i \rangle \\ &= \sum \langle [\mu_0^{-1} \zeta_a, \zeta_a], \zeta_b \rangle = d(\log \mathcal{N}) \zeta_b \end{aligned}$$

by (3.6). Similarly it is true that  $\langle \sum [\mu_0^{-1} \mathcal{A}^* \mathcal{A} v_i, \mathbb{I} v_i], \xi_\alpha \rangle = \sum \langle [\mu_0^{-1} \zeta_a, \zeta_a], \xi_\alpha \rangle$ . Using the property  $[Z_a, Y_\alpha] = \delta_{a,b} Z_c - \delta_{a,c} Z_b$  with  $Y_\alpha = [Z_b, Z_c]$  it is easy to see that

$$\sum \langle [\mu_0^{-1} \zeta_a, \zeta_a], \xi_\alpha \rangle = \langle \mu_0^{-1} \zeta_b, \zeta_c \rangle - \langle \mu_0^{-1} \zeta_c, \zeta_b \rangle = 0.$$

Thus  $\sum \mu_0^{-1} [\mu_0^{-1} \mathcal{A}^* \mathcal{A} v_i, \mathbb{I} v_i] = \mu_0^{-1} \sum (d(\log \mathcal{N}) \zeta_b) \zeta_b = \text{grad}^{\mu_0}(\log \mathcal{N})$  and claim (3.16) follows.

Claim:

$$(3.17) \quad \nabla_{\mu_0^{-1} \zeta_a}^{\text{nh}} (\mu_0^{-1} \zeta_a) = 0.$$

Again we use (3.14) to see that

$$\begin{aligned} \nabla_{\mu_0^{-1} \zeta_a}^{\text{nh}} (\mu_0^{-1} \zeta_a) &= \mu_0^{-1} [\mu_0^{-1} \zeta_a, \mathbb{I} \mu_0^{-1} \zeta_a] + \sum \langle \mathbb{I} v_j, \mu_0^{-1} \zeta_a \rangle (v_j \cdot \langle \mathbb{I} v_k, \mu_0^{-1} \zeta_a \rangle) v_k \\ &= \mu_0^{-1} [\mu_0^{-1} \zeta_a, \mathbb{I} \mu_0^{-1} \zeta_a] + \sum \langle \mathbb{I} v_j, \mu_0^{-1} \zeta_a \rangle \langle \mathbb{I} v_k, \mu_0^{-1} [v_j, \mathcal{A}^* \mathcal{A} \mu_0^{-1} \zeta_a] \rangle v_k \\ &\quad - \sum \langle \mathbb{I} v_j, \mu_0^{-1} \zeta_a \rangle \langle \mathbb{I} v_k, \mu_0^{-1} [v_j, \zeta_a] \rangle v_k \\ &= \mu_0^{-1} [\mu_0^{-1} \zeta_a, \mathbb{I} \mu_0^{-1} \zeta_a] + \mu_0^{-1} [\mu_0^{-1} \zeta_a, \mathcal{A}^* \mathcal{A} \mu_0^{-1} \zeta_a] - \mu_0^{-1} [\mu_0^{-1} \zeta_a, \zeta_a] \\ &= 0. \end{aligned}$$

Now (3.16) and (3.17) imply (3.15) which shows part (1) of the assertion.

For (2) one checks that (3.7) yields, for  $\xi_\alpha = [\zeta_c, \zeta_d]$ ,

$$\begin{aligned} \langle \text{grad}^{\mu_0}(\log \mathcal{N}), \xi_\alpha \rangle &= \sum \langle \mu_0^{-1} [\zeta_c, \zeta_d], \langle [\mu_0^{-1} \zeta_a, \zeta_a], \zeta_b \rangle \zeta_b \rangle \\ &= \sum \langle \mu_0^{-1} [\zeta_c, \zeta_d], \zeta_b \rangle \langle \mu_0^{-1} \zeta_a, [\zeta_a, \zeta_b] \rangle \\ &= \langle \mu_0^{-1} \zeta_d, [\zeta_c, \zeta_d] \rangle \langle \mu_0^{-1} \zeta_c, [\zeta_c, \zeta_d] \rangle (n-2) + \langle \mu_0^{-1} \zeta_c, [\zeta_c, \zeta_d] \rangle \langle \mu_0^{-1} \zeta_d, [\zeta_d, \zeta_c] \rangle (n-2) \\ &= 0 \end{aligned}$$

where we also have used that  $\sum \langle [\mu_0^{-1} \zeta_a, \zeta_a], \xi_\alpha \rangle = 0$ .  $\square$

*Proof of Corollary 3.4.* Part (1) is clear.

Concerning part (2) let  $f \in C^\infty(S^{n-1})$ . According to Definition 2.2 we should to show that the generator (3.11) satisfies

$$\left( \frac{1}{2} \sum (\mu_0^{-1} v_i) (\mu_0^{-1} v_i) + \frac{1}{2} \sum (\mu_0^{-1} \zeta_a) (\mu_0^{-1} \zeta_a) \right) \kappa^* f = \frac{1}{2} \kappa^* \Delta^\nu f$$

where  $\Delta^\nu$  is the Laplacian associated to  $\nu$ . Indeed, we find  $\frac{1}{2} \sum (\mu_0^{-1} v_i) (\mu_0^{-1} v_i) \kappa^* f + \frac{1}{2} \sum (\mu_0^{-1} \zeta_a) (\mu_0^{-1} \zeta_a) \kappa^* f = \frac{2}{2} \frac{1}{4} \sum \zeta_a \zeta_a \kappa^* f$ . Now  $\{\frac{1}{\sqrt{2}} \zeta_a\}$  is a horizontal orthonormal frame for  $\mu_0|(\text{Hor} \times \text{Hor})$  where  $\text{Hor}$  is the  $\mu_0$ -orthogonal to  $\ker T\kappa$ . Therefore,

$$\frac{1}{4} \sum \zeta_a \zeta_a \kappa^* f = \frac{1}{2} \text{Tr}^{\text{hor}} \text{Hess}^{\mu_0}(\kappa^* f) = \frac{1}{2} \kappa^* \Delta^\nu f$$

where  $\text{Tr}^{\text{hor}}$  denotes the trace computed with respect to horizontal fields only and  $\text{Hess}^{\mu_0}$  is the Hessian of the Levi-Civita connection on  $(K, \mu_0)$ . The equation  $\zeta_a \zeta_a \kappa^* f = \text{Hess}^{\mu_0}(\kappa^* f)(\zeta_a, \zeta_a)$  is justified by the observation that  $\nabla_{\zeta_a}^{\mu_0} \zeta_a = 0$  where  $\nabla^{\mu_0}$  is the Levi-Civita connection of the right invariant metric  $\mu_0$ . (In fact, the restriction of  $\nabla^{\text{nh}}$  to  $\text{Hor} \times \text{Hor}$  equals the restriction of  $\nabla^{\mu_0}$ .)  $\square$

In the homogeneous case the above construction yields a Brownian motion on  $S^{n-1}$  in a manner similar to the one described in [14, Chapter V] by the notion of rolling the sphere  $S^{n-1}$  along a Brownian motion in  $\mathbb{R}^{n-1}$  by means of the Levi-Civita connection. The difference is that [14] start from Brownian motion in  $\mathbb{R}^{n-1}$  while we started from Brownian motion in  $\mathbb{R}^m \times \mathbb{R}^{n-1}$  with  $m = \dim \mathfrak{so}(n) = \frac{n(n-1)}{2}$ . One can recover the setting of [14] by setting  $(H_0, H_1, \dots, H_m) = 0$  in (3.8). Then, with  $\mathbb{I} = 1$ , we obtain a diffusion  $\kappa \circ \tau \circ \Gamma$  which is driven by Brownian motion  $(W^1, \dots, W^{n-1})$  in  $\mathbb{R}^{n-1}$  and the generator of which is given by

$$\frac{1}{2} \sum (\mu_0^{-1} \zeta_a) (\mu_0^{-1} \zeta_a) \kappa^* f = \frac{1}{2} \kappa^* \Delta^\nu f$$

where  $f \in C^\infty(S^{n-1})$ . Referring to the interpretation stated in the introduction this means that the Chaplygin ball is subjected to horizontal jiggling but there is no angular jiggling. Equivalently, the ball sits on a table which undergoes a translational Brownian motion. Compare [19]. Alternatively, we can set  $(F_1, \dots, F_{n-1}) = 0$

in (3.8). Then there is only angular jiggling and the diffusion  $\tau \circ \Gamma$  is driven by  $(B^1, \dots, B^m)$ . By (3.17) the drift remains the same as in Theorem 3.3.

It seems that the notion of a stochastic non-holonomic system has been hardly investigated in the literature. We finish by asking the following questions.

- (1) Does an analog of Theorem 3.3 hold for general  $G$ -Chaplygin systems when there is a preserved measure? What can be said about the drift if there is no preserved measure?
- (2) Which is the precise relationship between  $G$ -Chaplygin systems with preserved measures and measure preserving ‘Chaplygin diffusions’? Preservation of measure by diffusions is studied in [14, Chapter V].
- (3) Is there a time change or Girsanov type argument to eliminate the drift in Theorem 3.3 or to make it even a Brownian motion? Is this related to the Hamiltonization of the deterministic problem?

#### 4. APPENDIX: $G$ -CHAPLYGIN SYSTEMS AND SYMMETRY REDUCTION

The purpose of this appendix is to shortly introduce and motivate the notion of a  $G$ -Chaplygin system and to state Proposition 4.1 which explains the symmetry reduction of such systems. This reduction is termed *compression* ([9]) to distinguish it from symplectic reduction. These concepts are closely related and compression can be viewed as a perturbed version of its symplectic counterpart. At the same time, however, there are fundamental differences; symmetries behave differently in non-holonomic mechanics and do not necessarily give rise to conserved quantities, and there need not exist a preserved measure ([7, Section 5.4]); all this is related to the question of closedness of the form  $\Omega_{\text{nh}}$  defined in Proposition 4.1. See [1, 3, 7, 9, 12].

A non-holonomic system is a triple  $(Q, \mathcal{D}, L)$  where  $Q$  is a configuration manifold,  $L : TQ \rightarrow \mathbb{R}$  is a Lagrangian, and  $\mathcal{D} \subset TQ$  is a smooth non-integrable distribution which is supposed to be of constant rank. The equations of motion for a curve  $q(t)$  which should satisfy  $q' \in \mathcal{D}$  are then stated in terms of the Lagrange d’Alembert principle. Suppose there is a Riemannian metric  $\mu$  on  $Q$  such that we have an isomorphism  $TQ \cong T^*Q$  and assume that  $L$  is the kinetic energy Lagrangian. In this case there is also an (almost) Hamiltonian version: continue to use the symbol  $\mu$  to denote the co-metric and consider the Hamiltonian  $\mathcal{H}(q, p)$  given by the Legendre transform of  $L$ . Since  $\mathcal{D}$  is of constant rank there is a family of independent one-forms  $\phi^a \in \Omega(Q)$  such that  $\mathcal{D}$  is the joint kernel of these. In terms of coordinates  $(q^i, p_i)$  the equations of motion are

$$(q^i)' = \frac{\partial \mathcal{H}}{\partial p_i} \text{ and } p_i' = -\frac{\partial \mathcal{H}}{\partial q^i} - \sum \lambda_a \phi^a \left( \frac{\partial}{\partial q^i} \right)$$

where the  $\lambda_a$  are the Lagrange multipliers to be determined from the supplementary condition that  $\mu(\phi^a, p) = 0$ . With  $X_{\mathcal{H}}^{\mathcal{C}} := (q', p')$  we may thus rephrase the equations as

$$(4.18) \quad i(X_{\mathcal{H}}^{\mathcal{C}})\Omega = d\mathcal{H} + \sum \lambda_a \tau^* \phi^a$$

where  $\Omega = -d\theta$  is the canonical symplectic form on  $T^*Q$  and  $\tau : T^*Q \rightarrow Q$  is the footpoint projection. (The notation  $X_{\mathcal{H}}^{\mathcal{C}}$  will become clear below.)

Let  $G$  be a Lie group that acts freely, properly and by isometries on the Riemannian manifold  $(Q, \mu)$ . A  $G$ -CHAPLYGIN SYSTEM is a non-holonomic system  $(Q, L = \frac{1}{2} \|\cdot\|_{\mu}^2, \mathcal{D})$  that has the property that  $\mathcal{D}$  is a principal connection on the principal bundle  $Q \twoheadrightarrow Q/G$ . Thus  $\mathcal{D}$  is the kernel of a connection form  $\mathcal{A} : TQ \rightarrow \mathfrak{g}$ . Notice that we do not require  $\mathcal{A}$  to be the mechanical connection associated to  $\mu$ .

Consider  $\mathcal{C}$  and  $\Omega^{\mathcal{C}}$  as defined in Section 2.C. Since  $X_{\mathcal{H}}^{\mathcal{C}}$  is, by construction, tangent to  $\mathcal{M}$  and takes values in  $\mathcal{C}$  one may now rewrite the equations of motion (4.18) in the appealing format

$$i(X_{\mathcal{H}}^{\mathcal{C}})\Omega^{\mathcal{C}} = (d\mathcal{H})^{\mathcal{C}}$$

where  $(d\mathcal{H})^{\mathcal{C}}$  is the restriction of  $\iota^*d\mathcal{H}$  to  $\mathcal{C}$  with  $\iota : \mathcal{D} \hookrightarrow TQ$  being the inclusion.

Let  $\mu_0$  denote the induced metric on  $S := Q/G$  that makes  $\pi : Q \twoheadrightarrow S$  a Riemannian submersion. Identify tangent and cotangent space of  $Q$  and  $S$  via their respective metrics. Consider the orbit projection map

$$\rho := T\pi|_{\mathcal{D}} : \mathcal{D} \twoheadrightarrow \mathcal{D}/G = TS.$$

We may also associate a fiber-wise inverse to this mapping which is given by the horizontal lift mapping  $\text{hl}^A$  associated to  $\mathcal{A}$ . As already noted in Section 2.C,  $\iota^*\tau^*\mathcal{A} : T\mathcal{D} \rightarrow \mathfrak{g}$  defines a principal bundle connection for  $\rho$ , whose horizontal space is given by  $\mathcal{C}$ .

**Proposition 4.1** (Compression). *The following are true.*

- (1)  $\Omega^{\mathcal{C}}$  descends to a non-degenerate two-form  $\Omega_{\text{nh}}$  on  $TS$ .
- (2)  $\Omega_{\text{nh}} = \Omega_S - \langle J_G \circ \text{hl}^A, \tau_S^* \text{Curv}^A \rangle$ . Here  $\Omega_S = -d\theta_S$  is the canonical form on  $TS$ ,  $J_G$  is the momentum map of the tangent lifted  $G$ -action on  $TQ$ ,  $\text{Curv}^A \in \Omega^2(S, \mathfrak{g})$  is the curvature form of  $\mathcal{A}$ , and  $\tau_S : TS \rightarrow S$  is the projection.
- (3) Let  $h : TQ \rightarrow \mathbb{R}$  be  $G$ -invariant. Then the vectorfield

$$X_h^{\mathcal{C}} := (\Omega^{\mathcal{C}})^{-1}(dh)^{\mathcal{C}},$$

where  $(dh)^{\mathcal{C}}$  is the restriction of  $\iota^*dh$  to  $\mathcal{C}$ , is  $\rho$ -related to the vector field  $X_{h_0}^{\text{nh}}$  on  $T^*S$  defined by

$$i(X_{h_0}^{\text{nh}})\Omega_{\text{nh}} = dh_0$$

where the compressed Hamiltonian,  $h_0 : T^*S = TS \rightarrow \mathbb{R}$  is defined by  $h_0 := h \circ \iota \circ \text{hl}^A$ , with  $\text{hl}^A$  denoting the horizontal lift mapping.

In general,  $\Omega_{\text{nh}}$  is an almost symplectic form, that is, it is non-degenerate and non-closed. See [1, 9, 12].

#### REFERENCES

- [1] L. Bates and J. Sniatycki, *Nonholonomic reduction*, Rep. Math. Phys. **32**, No. 1, 99-115, 1993.
- [2] J.-M. Bismut, *Mecanique Aleatoire*, LN in Math., vol. 866, Springer 1981.
- [3] A.M. Bloch, *Nonholonomic mechanics and control*, Springer, 2003.
- [4] A.V. Borisov and I.S. Mamaev, *Chaplygin's ball rolling problem is Hamiltonian*, Mathematical Notes, **70** (2001), 793-795.
- [5] A.V. Borisov and I.S. Mamaev, *Hamiltonization of nonholonomic systems*, (2005), [arXiv:nlin/0509036v1](https://arxiv.org/abs/nlin/0509036v1).
- [6] S.A. Chaplygin, *On a ball's rolling on a horizontal plane*, Regul. Chaotic Dyn., **7** (2002), 131-148; Translation of original in: Mathematical collection of the Moscow Mathematical Society, **24** (1903), 139-168, (Russian).
- [7] J. Cortes Monforte, *Geometric control and numerical aspects of non-holonomic systems*, LN in Math 1793, Springer, 2002.
- [8] J.J. Duistermaat, *Chaplygin's sphere*, [arXiv:math/0409019v1](https://arxiv.org/abs/math/0409019v1).
- [9] K. Ehlers, J. Koiller, R. Montgomery and P.M. Rios, *Nonholonomic systems via moving frames: Cartan equivalence and Chaplygin Hamiltonization*, in "The breath of Symplectic and Poisson Geometry," Progress in Mathematics, **232** (2004), 75-120.
- [10] M. Emery, *Stochastic calculus in manifolds*, Universitext, Springer, 1989.
- [11] Y.N. Fedorov and V.V. Kozlov, *Various aspects of n-dimensional rigid body dynamics*, Amer. Math. Soc. Transl., Ser. (2), **168** (1995), 141-171.
- [12] S. Hochgerner and L. Garcia-Naranjo, *G-Chaplygin systems with internal symmetries, Truncation, and an (almost) symplectic view of Chaplygin's ball*, J. Geom. Mech. **1**, No. 1, pp. 35-53, 2009.
- [13] S. Hochgerner, *Chaplygin systems associated to Cartan decompositions of semi-simple Lie groups*, [arXiv:0907.0636](https://arxiv.org/abs/0907.0636).
- [14] N. Ikeda, S. Watanabe, *Stochastic differential equations and diffusion processes*, North-Holland Publishing Company, Kodansha, 2nd Ed., 1989.
- [15] B. Jovanovic, *LR and L+R systems*, J. Phys. A: Math. Theor. **42** No. 22, (2009).
- [16] B. Jovanovic, *Hamiltonization and integrability of the Chaplygin sphere in  $\mathbb{R}^n$* , [arXiv:math-ph/0902.4397v1](https://arxiv.org/abs/math-ph/0902.4397v1).
- [17] J.-A. Lazaro-Cami, J.-P. Ortega, *Stochastic Hamiltonian dynamical systems*, Rep. Math. Phys. **61** (2008), pp. 65-122.
- [18] ———, *Reduction, reconstruction, and skew-product decomposition of symmetric stochastic differential equations*, [arXiv:0705.3156v2](https://arxiv.org/abs/0705.3156v2), 2008.
- [19] N.K. Moshchuk, I.N. Sinitsyn, *On stochastic non-holonomic systems*, PMM USSR, Vol. **54**, No. 2 (1990), pp. 174-182.
- [20] B. Oksendal, *Stochastic differential equations*, Universitext, Springer, 2007.

SECTION DE MATHÉMATIQUES, STATION 8, EPFL, CH-1015 LAUSANNE

E-mail address: [simon.hochgerner@epfl.ch](mailto:simon.hochgerner@epfl.ch)